ENTERED

OIPE

RAW SEQUENCE LISTING DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:58

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

```
 3 <110> APPLICANT: Watowich, Stanley J.
 4       Weaver, Scott C.
 5       Davey, Robert A.
 7 <120> TITLE OF INVENTION: Drug Discovery Methods
 9 <130> FILE REFERENCE: 265.00260101
11 <140> CURRENT APPLICATION NUMBER: US 09/981,286A
12 <141> CURRENT FILING DATE: 2001-10-15
14 <150> PRIOR APPLICATION NUMBER: US 60/240,187
15 <151> PRIOR FILING DATE: 2000-10-13
17 <160> NUMBER OF SEQ ID NOS: 36
19 <170> SOFTWARE: PatentIn version 3.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 157
23 <212> TYPE: PRT
24 <213> ORGANISM: VENEZUELAN EQUINE ENCEPHALITIS VIRUS
26 <400> SEQUENCE: 1
28 Val Met Lys Leu Glu Ser Asp Lys Thr Phe Pro Ile Met Leu Glu Gly
29 1               5                  10                  15
31 Lys Ile Asn Gly Tyr Ala Cys Val Val Gly Gly Lys Leu Phe Arg Pro
32             20                  25                  30
34 Met His Val Glu Gly Lys Ile Asp Asn Asp Val Leu Ala Ala Leu Lys
35         35                  40                  45
37 Thr Lys Lys Ala Ser Lys Tyr Asp Leu Glu Tyr Ala Asp Val Pro Gln
38     50                  55                  60
40 Asn Met Arg Ala Asp Thr Phe Lys Tyr Thr His Glu Lys Pro Gln Gly
41 65                  70                  75                  80
43 Tyr Tyr Ser Trp His His Gly Ala Val Gln Tyr Glu Asn Gly Arg Phe
44                 85                  90                  95
46 Thr Val Pro Lys Gly Val Gly Ala Lys Gly Asp Ser Gly Arg Pro Ile
47             100                 105                 110
49 Leu Asp Asn Gln Gly Arg Val Val Ala Ile Val Leu Gly Gly Val Asn
50         115                 120                 125
52 Glu Gly Ser Arg Thr Ala Leu Ser Val Val Met Trp Asn Glu Lys Gly
53     130                 135                 140
55 Val Thr Val Lys Tyr Thr Pro Glu Asn Cys Glu Gln Trp
56 145                 150                 155
58 <210> SEQ ID NO: 2
59 <211> LENGTH: 11
60 <212> TYPE: PRT
C--> 61 <213> ORGANISM: ARTIFICIAL
63 <220> FEATURE:
64 <223> OTHER INFORMATION: Cell-permeant polypeptide
66 <400> SEQUENCE: 2
```

RAW SEQUENCE LISTING DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:58

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

```
     68 Tyr Gly Arg Lys Lys Arg Arg Gln Arg Arg Arg
     69 1               5                   10
     71 <210> SEQ ID NO: 3
     72 <211> LENGTH: 16
     73 <212> TYPE: PRT
C--> 74 <213> ORGANISM: ARTIFICIAL
     76 <220> FEATURE:
     77 <223> OTHER INFORMATION: Cell-permeant polypeptide
     79 <400> SEQUENCE: 3
     81 Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys
     82 1               5                   10                  15
     84 <210> SEQ ID NO: 4
     85 <211> LENGTH: 16
     86 <212> TYPE: PRT
C--> 87 <213> ORGANISM: ARTIFICIAL
     89 <220> FEATURE:
     90 <223> OTHER INFORMATION: Cell-permeant polypeptide
     92 <400> SEQUENCE: 4
     94 Arg Gln Ile Lys Ile Trp Phe Pro Asn Arg Arg Met Lys Trp Lys Lys
     95 1               5                   10                  15
     97 <210> SEQ ID NO: 5
     98 <211> LENGTH: 16
     99 <212> TYPE: PRT
C--> 100 <213> ORGANISM: ARTIFICIAL
    102 <220> FEATURE:
    103 <223> OTHER INFORMATION: Cell-permeant polypeptide
    105 <400> SEQUENCE: 5
    107 Arg Gln Pro Lys Ile Trp Phe Pro Asn Arg Arg Pro Lys Trp Lys Lys
    108 1               5                   10                  15
    110 <210> SEQ ID NO: 6
    111 <211> LENGTH: 525
    112 <212> TYPE: DNA
C--> 113 <213> ORGANISM: ARTIFICIAL
    115 <220> FEATURE:
    116 <223> OTHER INFORMATION: Nucleotide sequence encoding tat-CCD
    118 <400> SEQUENCE: 6
    119 atgtacggtc gtaaaaaacg tcgtcagcgt cgtcgtgtca tgaaattgga atctgacaag     60
    121 acgttcccaa tcatgttgga agggaagata aacggctacg cttgtgtggt cggagggaag    120
    123 ttattcaggc cgatgcatgt ggaaggcaag atcgacaacg acgttctggc cgcgcttaag    180
    125 acgaagaaag catccaaata cgatcttgag tatgcagatg tgccacagaa catgcgggcc    240
    127 gatacattca aatacaccca tgagaaaccc caaggctatt acagctggca tcatggagca    300
    129 gtccaatatg aaaatgggcg tttcacggtg ccgaaaggag ttggggccaa gggagacagc    360
    131 ggacgaccca ttctggataa ccagggacgg gtggtcgcta ttgtgctggg aggtgtgaat    420
    133 gaaggatcta ggacagccct ttcagtcgtc atgtggaaca agcttggatc ttctctcgag    480
    135 ggagttaccg tgaagtatac tccggagaac tgcgagcaat ggtaa                    525
    138 <210> SEQ ID NO: 7
    139 <211> LENGTH: 169
    140 <212> TYPE: PRT
C--> 141 <213> ORGANISM: ARTIFICIAL
```

RAW SEQUENCE LISTING DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:58

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

<220> FEATURE:
<223> OTHER INFORMATION: Amino acid sequence of tat-CCD
<400> SEQUENCE: 7

Met	Tyr	Gly	Arg	Lys	Lys	Arg	Arg	Gln	Arg	Arg	Arg	Val	Met	Lys	Leu
1				5				10						15	
Glu	Ser	Asp	Lys	Thr	Phe	Pro	Ile	Met	Leu	Glu	Gly	Lys	Ile	Asn	Gly
		20					25						30		
Tyr	Ala	Cys	Val	Val	Gly	Gly	Lys	Leu	Phe	Arg	Pro	Met	His	Val	Glu
	35					40					45				
Gly	Lys	Ile	Asp	Asn	Asp	Val	Leu	Ala	Ala	Leu	Lys	Thr	Lys	Lys	Ala
	50				55					60					
Ser	Lys	Tyr	Asp	Leu	Glu	Tyr	Ala	Asp	Val	Pro	Gln	Asn	Met	Arg	Ala
65				70					75						80
Asp	Thr	Phe	Lys	Tyr	Thr	His	Glu	Lys	Pro	Gln	Gly	Tyr	Tyr	Ser	Trp
			85					90					95		
His	His	Gly	Ala	Val	Gln	Tyr	Glu	Asn	Gly	Arg	Phe	Thr	Val	Pro	Lys
		100					105					110			
Gly	Val	Gly	Ala	Lys	Gly	Asp	Ser	Gly	Arg	Pro	Ile	Leu	Asp	Asn	Gln
	115					120					125				
Gly	Arg	Val	Val	Ala	Ile	Val	Leu	Gly	Gly	Val	Asn	Glu	Gly	Ser	Arg
	130				135					140					
Thr	Ala	Leu	Ser	Val	Val	Met	Trp	Asn	Glu	Lys	Gly	Val	Thr	Val	Lys
145				150					155						160
Tyr	Thr	Pro	Glu	Asn	Cys	Glu	Gln	Trp							
			165												

<210> SEQ ID NO: 8
<211> LENGTH: 124
<212> TYPE: PRT
<213> ORGANISM: BOS TAURUS
<400> SEQUENCE: 8

Lys	Glu	Thr	Ala	Ala	Ala	Lys	Phe	Glu	Arg	Gln	His	Met	Asp	Ser	Ser
1				5				10						15	
Thr	Ser	Ala	Ala	Ser	Ser	Ser	Asn	Tyr	Cys	Asn	Gln	Met	Met	Lys	Ser
		20					25						30		
Arg	Asn	Leu	Thr	Lys	Asp	Arg	Cys	Lys	Pro	Val	Asn	Thr	Phe	Val	His
	35					40					45				
Glu	Ser	Leu	Ala	Asp	Val	Gln	Ala	Val	Cys	Ser	Gln	Lys	Asn	Val	Ala
	50				55					60					
Cys	Lys	Asn	Gly	Gln	Thr	Asn	Cys	Tyr	Gln	Ser	Tyr	Ser	Thr	Met	Ser
65				70					75						80
Ile	Thr	Asp	Cys	Arg	Glu	Thr	Gly	Ser	Ser	Lys	Tyr	Pro	Asn	Cys	Ala
			85					90					95		
Tyr	Lys	Thr	Thr	Gln	Ala	Asn	Lys	His	Ile	Ile	Val	Ala	Cys	Glu	Gly
		100					105					110			
Asn	Pro	Tyr	Val	Pro	Val	His	Phe	Ala	Ala	Ser	Val				
	115					120									

<210> SEQ ID NO: 9
<211> LENGTH: 37
<212> TYPE: DNA

RAW SEQUENCE LISTING DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:58

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

C--> 215 <213> ORGANISM: ARTIFICIAL
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Primer
220 <400> SEQUENCE: 9
221 gggaattcca tatggtcatg aattggaatc tgacaag 37
224 <210> SEQ ID NO: 10
225 <211> LENGTH: 42
226 <212> TYPE: DNA
C--> 227 <213> ORGANISM: ARTIFICIAL
229 <220> FEATURE:
230 <223> OTHER INFORMATION: Primer
232 <400> SEQUENCE: 10
233 gaattcggat cctcattacc attgctcgca gttctccgga gt 42
236 <210> SEQ ID NO: 11
237 <211> LENGTH: 6
238 <212> TYPE: PRT
C--> 239 <213> ORGANISM: ARTIFICIAL
241 <220> FEATURE:
242 <223> OTHER INFORMATION: A variable region amino acid sequence
244 <220> FEATURE:
245 <221> NAME/KEY: Variant
246 <222> LOCATION: (1)..(6)
247 <223> OTHER INFORMATION: Any amino acid
250 <400> SEQUENCE: 11
W--> 252 Xaa Xaa Xaa Xaa Xaa Xaa
253 1 5
255 <210> SEQ ID NO: 12
256 <211> LENGTH: 477
257 <212> TYPE: DNA
258 <213> ORGANISM: VENEZUELAN EQUINE ENCEPHALITIS VIRUS
260 <400> SEQUENCE: 12
261 gtcatgaaat tggaatctga caagacgttc ccaatcatgt tggaagggaa gataaacggc 60
263 tacgcttgtg tggtcggagg gaagttattc aggccgatgc atgtggaagg caagatcgac 120
265 aacgacgttc tggccgcgct taagacgaag aaagcatcca aatacgatct tgagtatgca 180
267 gatgtgccac agaacatgcg ggccgataca ttcaaataca cccatgagaa accccaaggc 240
269 tattacagct ggcatcatgg agcagtccaa tatgaaaatg ggcgtttcac ggtgccgaaa 300
271 ggagttgggg ccaagggaga cagcggacga cccattctgg ataaccaggg acgggtggtc 360
273 gctattgtgc tgggaggtgt gaatgaagga tctaggacag ccctttcagt cgtcatgtgg 420
275 aacgagaagg gagttaccgt gaagtatact ccggagaact gcgagcaatg gtaatga 477
278 <210> SEQ ID NO: 13
279 <211> LENGTH: 43
280 <212> TYPE: DNA
C--> 281 <213> ORGANISM: ARTIFICIAL
283 <220> FEATURE:
284 <223> OTHER INFORMATION: Primer
286 <400> SEQUENCE: 13
287 agctaggaat tcggatccca tatgtacggt cgtaaaaaac gtc 43
290 <210> SEQ ID NO: 14
291 <211> LENGTH: 33

RAW SEQUENCE LISTING DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:58

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

292 <212> TYPE: DNA
C--> 293 <213> ORGANISM: ARTIFICIAL
295 <220> FEATURE:
296 <223> OTHER INFORMATION: Primer
298 <400> SEQUENCE: 14
299 ctagctaagc ttgttccaca tgacgactga aag 33
302 <210> SEQ ID NO: 15
303 <211> LENGTH: 36
304 <212> TYPE: DNA
C--> 305 <213> ORGANISM: ARTIFICIAL
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Primer
310 <400> SEQUENCE: 15
311 ctagctgcgg ccgctcatta ccattgctcg cagttc 36
314 <210> SEQ ID NO: 16
315 <211> LENGTH: 47
316 <212> TYPE: DNA
C--> 317 <213> ORGANISM: ARTIFICIAL
319 <220> FEATURE:
320 <223> OTHER INFORMATION: Primer
322 <400> SEQUENCE: 16
323 agctagaagc ttggatcttc tctcgaggga gttaccgtga agtatac 47
326 <210> SEQ ID NO: 17
327 <211> LENGTH: 50
328 <212> TYPE: DNA
C--> 329 <213> ORGANISM: ARTIFICIAL
331 <220> FEATURE:
332 <223> OTHER INFORMATION: Primer
334 <400> SEQUENCE: 17
335 gatcctcgag agaagatccg gatccgttcc acatgacgac tgaaagggct 50
338 <210> SEQ ID NO: 18
339 <211> LENGTH: 51
340 <212> TYPE: DNA
C--> 341 <213> ORGANISM: ARTIFICIAL
343 <220> FEATURE:
344 <223> OTHER INFORMATION: Primer
346 <400> SEQUENCE: 18
347 gatcgaattc caccagcaga atcgacatat gtacggtcgt aaaaaacgtc g 51
350 <210> SEQ ID NO: 19
351 <211> LENGTH: 27
352 <212> TYPE: DNA
C--> 353 <213> ORGANISM: ARTIFICIAL
355 <220> FEATURE:
356 <223> OTHER INFORMATION: Primer
358 <220> FEATURE:
359 <221> NAME/KEY: misc_feature
360 <222> LOCATION: (15)..(16)
361 <223> OTHER INFORMATION: A, T, G, or C
364 <400> SEQUENCE: 19

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/28/2002
PATENT APPLICATION: US/09/981,286A TIME: 14:15:59

Input Set : A:\26500260101.ST25.txt
Output Set: N:\CRF3\06282002\I981286A.raw

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:11; Xaa Pos. 1,2,3,4,5,6
Seq#:19; N Pos. 15,16
Seq#:23; N Pos. 15,16
Seq#:25; N Pos. 16,17
Seq#:26; N Pos. 20,21
Seq#:28; N Pos. 19,20,22,23,25,26,28,29,31,32,34,35
Seq#:30; N Pos. 1,2,4,5,7,8,10,11,13,14,16,17

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:2,3,4,5,6,7,9,10,11,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30
Seq#:31,32,33,34,35,36

VERIFICATION SUMMARY DATE: 06/28/2002
PATENT APPLICATION: **US/09/981,286A** TIME: 14:15:59

Input Set : **A:\26500260101.ST25.txt**
Output Set: **N:\CRF3\06282002\I981286A.raw**

L:61 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:2
L:74 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:3
L:87 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:4
L:100 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:5
L:113 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:6
L:141 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:7
L:215 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:9
L:227 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:10
L:239 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:11
L:252 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:0
L:281 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:13
L:293 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14
L:305 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15
L:317 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16
L:329 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17
L:341 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:18
L:353 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:19
L:365 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19 after pos.:0
L:371 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:20
L:383 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21
L:395 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22
L:407 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:23
L:419 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23 after pos.:0
L:425 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:24
L:437 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:25
L:449 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25 after pos.:0
L:455 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:26
L:467 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26 after pos.:0
L:473 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:27
L:485 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:28
L:527 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28 after pos.:0
L:533 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:29
L:545 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:30
L:587 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30 after pos.:0
L:593 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:31
L:621 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:32
L:649 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:33
L:692 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:34
L:735 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:35
L:747 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:36